second sheet, and the long helix breaks and becomes distorted, as actually occurs in 2pcy's native structure. Most of the displaced residues joined the tube fragments generated by various secondary structural elements of the template, but only a few maintained their original assignments to the template tube. This way the internal force field of the lattice model overrode the target interactions, significantly correcting the threading model. The initial model and the final model are compared with the native structure in Figure 16, where stereo alpha-carbon traces are displayed in their best mutual superposition, using the MOLMOL[20] drawing program.

256bA

With regard to this protein, there was a four-helix bundle and the threading alignment had a few gaps. The template structure was very similar to the target, but the threading model was not very good. During the simulations, most of the C-terminal helical hairpin remained almost unchanged, except for the loop region that was very mobile. The third (first helix of the C-terminal hairpin) helix of the model was the most stable. The N-terminal hairpin underwent a large-scale rearrangement. The second helix underwent a rotation that changed its packing angle with respect to the remainder of the molecule. As a result, the end of this helix moved by about 7 Å in a lateral direction, while the beginning of this helix stayed close to its original position. The largest changes were observed for the first N-terminal helix. It moved along the tube, changing assignment indices by several residues (up to 8); a lateral adjustment took place as well. The initial model and the final model (superimposed onto the native structure) are compared in Figure 17. The helical regions of the final model are very close to the native structure; the largest errors that account for most of the structure errors are in the central turn/loop region.

1tlk

Telokin is a quite regular β-protein. Again, due to gaps and insertions, the threading model for it produced a wrong topology. During the simulations, one of the β-strands from the original model left the initial assignment and stuck to the tube of a strand from the opposite sheet. Two β-strands that were not in the threading model (lack of the alignment assignments) were built in the simulated annealing procedure, and they joined tubes associated with existing strands. The entire structure, except for the last β-terminal β-strand that remained essentially unchanged, rearranged substantially. Mostly lateral (orthogonal to the local direction of the template tube) displacements occurred in the range of 6 Å for about half of all the residues. As a result, the model improved its RMSD by almost 4 Å. The initial model and the final model (superimposed onto the native structure) are compared in Figure 18.

How to identify good models

As mentioned above, the instant invention generates low to moderate resolution models of correct topology in those cases when the initial threading-based alignment leads to at least a partially correct structure, *i.e.*, where a part of the identified template is close to the target structure. How to (*a priori*) distinguish a good (threading-based) alignment from a poor one is a non-trivial question. Unfortunately, there is not yet a general solution to this problem.

The intrinsic force field of the reduced model correctly identifies the native structure (the lattice protection) as the lowest energy conformation when compared with the models generated by MODELLER from the initial threading alignments. The models obtained in the lattice homology modeling are described herein. In all cases except one (lbbhA, where MODELLER gave a slightly better result than the present method), the energy of the models built by the present method is significantly lower than other worse models (including these built by automatic use of MODELLER). While interesting, there remains a need to be able to distinguish

those target/template pairs where the final model is of reasonable quality from those cases where, despite a sometimes large improvement of the initial models, the resulting structures are still far from the native target conformation. Unfortunately, simple energetic criteria (conformational energy per residue in the final model, decrease of energy from the starting model to the final model, *etc.*) do not enable identification of these poor quality structures.

The previous section discussed how the modeling procedure of the invention improves the initial, threading-based model. This could be actually used for a qualitative identification of better models. Consider the displacement of particular residues (as a function of their position along the chain) during the entire simulation procedure. In those cases where the final model is of good quality, the plots indicate relatively well separated regions where the chain modifications were small and also indicated regions of large modifications. This is consistent with the previously mentioned characteristic behavior of "good" models, for which some ligaments of alignments are recognized by the procedure as being very good and behave as a scaffold for readjustment of the remainder of the protein. In contrast, poor models are characterized by random fluctuations of the spatial amino acid displacements along the sequence. In such cases there is no pattern. Perhaps, there is a huge energy barrier between the starting model and the better, near native models that cannot be surmounted by partial readjustment of the initial alignment. Examples of both situations are given in Figures 19 and 20. The lowest (and locally similar) displacement (during the modeling procedure) regions identify the regions of an optimal (or very close to optimal) alignment. While the above is not easy for a simple quantification, it still can be used as a heuristic criterion for the identification of cases where the method proposed in this work is likely to provide relatively good, low resolution models. Figure 21 shows the plot of model accuracy (measured as the alpha carbon RMSD from native) as a function of the variability in the model chain mobility during the simulations. Unfortunately, the correlation is not very strong. Consequently, the mobility criterion has to be used with caution. Rather